

Robinson



1600

RAW SEQUENCE LISTING

DATE: 02/11/2003

PATENT APPLICATION: US/09/717,095

TIME: 20:11:15

Input Set : N:\Crf3\RULE60\09717095.txt

Output Set: N:\CRF4\02112003\I717095.raw

SEQUENCE LISTING

ENTERED

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Vinik, Aaron I.

6 Pittenger, Gary L.

7 Rafaeloff, Ronit

8 Rosenberg, Lawrence

9 Duguid, William P.

12 (ii) TITLE OF INVENTION: INGAP PROTEIN INVOLVED IN PANCREATIC
13 ISLET NEOGENESIS

15 (iii) NUMBER OF SEQUENCES: 7

17 (iv) CORRESPONDENCE ADDRESS:

18 (A) ADDRESSEE: Banner & Witcoff, Ltd.

19 (B) STREET: 1001 G Street, N.W.

20 (C) CITY: Washington

21 (D) STATE: D.C.

22 (E) COUNTRY: US

23 (F) ZIP: 20001-4597

25 (v) COMPUTER READABLE FORM:

26 (A) MEDIUM TYPE: Floppy disk

27 (B) COMPUTER: IBM PC compatible

28 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

29 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

31 (vi) CURRENT APPLICATION DATA:

C--> 32 (A) APPLICATION NUMBER: US/09/717,095

C--> 33 (B) FILING DATE: 22-Nov-2000

34 (C) CLASSIFICATION:

36 (vii) PRIOR APPLICATION DATA:

37 (A) APPLICATION NUMBER: US/08/709,662

38 (B) FILING DATE: 09-SEP-1996

40 (viii) ATTORNEY/AGENT INFORMATION:

41 (A) NAME: Kagan, Sarah A.

42 (B) REGISTRATION NUMBER: 32,141

43 (C) REFERENCE/DOCKET NUMBER: 00570.59178

45 (ix) TELECOMMUNICATION INFORMATION:

46 (A) TELEPHONE: 202-508-9100

47 (B) TELEFAX: 202-508-9299

50 (2) INFORMATION FOR SEQ ID NO: 1:

52 (i) SEQUENCE CHARACTERISTICS:

53 (A) LENGTH: 747 base pairs

54 (B) TYPE: nucleic acid

55 (C) STRANDEDNESS: single

56 (D) TOPOLOGY: linear

58 (ii) MOLECULE TYPE: cDNA

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/717,095

DATE: 02/11/2003

TIME: 20:11:15

Input Set : N:\Crif3\RULE60\09717095.txt

Output Set: N:\CRF4\02112003\I717095.raw

```

60      (iii) HYPOTHETICAL: NO
62      (iv) ANTI-SENSE: NO
64      (vi) ORIGINAL SOURCE:
65          (A) ORGANISM: Cricetulus
67      (ix) FEATURE:
68          (A) NAME/KEY: CDS
69          (B) LOCATION: 20..541
72      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
74 CTGCAAGACA GGTACCATG ATG CTT CCC ATG ACC CTC TGT AGG ATG TCT TGG      52
75          Met Leu Pro Met Thr Leu Cys Arg Met Ser Trp
76              1              5              10
78 ATG CTG CTT TCC TGC CTG ATG TTC CTT TCT TGG GTG GAA GGT GAA GAA      100
79 Met Leu Leu Ser Cys Leu Met Phe Leu Ser Trp Val Glu Gly Glu Glu
80              15              20              25
82 TCT CAA AAG AAA CTG CCT TCT TCA CGT ATA ACC TGT CCT CAA GGC TCT      148
83 Ser Gln Lys Lys Leu Pro Ser Ser Arg Ile Thr Cys Pro Gln Gly Ser
84              30              35              40
86 GTA GCC TAT GGG TCC TAT TGC TAT TCA CTG ATT TTG ATA CCA CAG ACC      196
87 Val Ala Tyr Gly Ser Tyr Cys Tyr Ser Leu Ile Leu Ile Pro Gln Thr
88              45              50              55
90 TGG TCT AAT GCA GAA CTA TCC TGC CAG ATG CAT TTC TCA GGA CAC CTG      244
91 Trp Ser Asn Ala Glu Leu Ser Cys Gln Met His Phe Ser Gly His Leu
92 60              65              70              75
94 GCA TTT CTT CTC AGT ACT GGT GAA ATT ACC TTC GTG TCC TCC CTT GTG      292
95 Ala Phe Leu Leu Thr Gly Glu Ile Thr Phe Val Ser Ser Leu Val
96              80              85              90
98 AAG AAC AGT TTG ACG GCC TAC CAG TAC ATC TGG ATT GGA CTC CAT GAT      340
99 Lys Asn Ser Leu Thr Ala Tyr Gln Tyr Ile Trp Ile Gly Leu His Asp
100              95              100              105
102 CCC TCA CAT GGT ACA CTA CCC AAC GGA AGT GGA TGG AAG TGG AGC AGT      388
103 Pro Ser His Gly Thr Leu Pro Asn Gly Ser Gly Trp Lys Trp Ser Ser
104              110              115              120
106 TCC AAT GTG CTG ACC TTC TAT AAC TGG GAG AGG AAC CCC TCT ATT GCT      436
107 Ser Asn Val Leu Thr Phe Tyr Asn Trp Glu Arg Asn Pro Ser Ile Ala
108              125              130              135
110 GCT GAC CGT GGT TAT TGT GCA GTT TTG TCT CAG AAA TCA GGT TTT CAG      484
111 Ala Asp Arg Gly Tyr Cys Ala Val Leu Ser Gln Lys Ser Gly Phe Gln
112 140              145              150              155
114 AAG TGG AGA GAT TTT AAT TGT GAA AAT GAG CTT CCC TAT ATC TGC AAA      532
115 Lys Trp Arg Asp Phe Asn Cys Glu Asn Glu Leu Pro Tyr Ile Cys Lys
116              160              165              170
118 TTC AAG GTC TAGGGCAGTT CTAATTTCAA CAGCTTGAAA ATATTATGAA      581
119 Phe Lys Val
122 GCTCACATGG ACAAGGAAGC AAGTATGAGG ATTCACTCAG GAAGAGCAAG CTCTGCCTAC      641
124 ACACCCACAC GAATTCCTT ATATCATCTC TGCTGTTTTT CTATCAGTAT ATTCTGTGGT      701
126 GGCTGTAACC TAAAGGCTCA GAGAACAAAA ATAAAATGTC ATCAAC      747
129 (2) INFORMATION FOR SEQ ID NO: 2:
131      (i) SEQUENCE CHARACTERISTICS:
132          (A) LENGTH: 174 amino acids

```

RAW SEQUENCE LISTING

DATE: 02/11/2003

PATENT APPLICATION: US/09/717,095

TIME: 20:11:15

Input Set : N:\Crif3\RULE60\09717095.txt

Output Set: N:\CRF4\02112003\I717095.raw

```

133          (B) TYPE: amino acid
134          (D) TOPOLOGY: linear
136          (ii) MOLECULE TYPE: protein
138          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
140 Met Leu Pro Met Thr Leu Cys Arg Met Ser Trp Met Leu Leu Ser Cys
141   1      5      10      15
143 Leu Met Phe Leu Ser Trp Val Glu Gly Glu Ser Gln Lys Lys Leu
144      20      25      30
146 Pro Ser Ser Arg Ile Thr Cys Pro Gln Gly Ser Val Ala Tyr Gly Ser
147      35      40      45
149 Tyr Cys Tyr Ser Leu Ile Leu Ile Pro Gln Thr Trp Ser Asn Ala Glu
150      50      55      60
152 Leu Ser Cys Gln Met His Phe Ser Gly His Leu Ala Phe Leu Leu Ser
153   65      70      75      80
155 Thr Gly Glu Ile Thr Phe Val Ser Ser Leu Val Lys Asn Ser Leu Thr
156      85      90      95
158 Ala Tyr Gln Tyr Ile Trp Ile Gly Leu His Asp Pro Ser His Gly Thr
159      100     105     110
161 Leu Pro Asn Gly Ser Gly Trp Lys Trp Ser Ser Ser Asn Val Leu Thr
162      115     120     125
164 Phe Tyr Asn Trp Glu Arg Asn Pro Ser Ile Ala Ala Asp Arg Gly Tyr
165      130     135     140
167 Cys Ala Val Leu Ser Gln Lys Ser Gly Phe Gln Lys Trp Arg Asp Phe
168 145      150     155     160
170 Asn Cys Glu Asn Glu Leu Pro Tyr Ile Cys Lys Phe Lys Val
171      165     170
173 (2) INFORMATION FOR SEQ ID NO: 3:
175   (i) SEQUENCE CHARACTERISTICS:
176       (A) LENGTH: 175 amino acids
177       (B) TYPE: amino acid
178       (D) TOPOLOGY: linear
180   (ii) MOLECULE TYPE: protein
182   (iv) ANTI-SENSE: NO
184   (vi) ORIGINAL SOURCE:
185       (A) ORGANISM: Rattus rattus
188   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
190 Met Leu His Arg Leu Ala Phe Pro Val Met Ser Trp Met Leu Leu Ser
191   1      5      10      15
193 Cys Leu Met Leu Leu Ser Gln Val Gln Gly Glu Asp Ser Pro Lys Lys
194      20      25      30
196 Ile Pro Ser Ala Arg Ile Ser Cys Pro Lys Gly Ser Gln Ala Tyr Gly
197      35      40      45
199 Ser Tyr Cys Tyr Ala Leu Phe Gln Ile Pro Gln Thr Trp Phe Asp Ala
200      50      55      60
202 Glu Leu Ala Cys Gln Lys Arg Pro Glu Gly His Leu Val Ser Val Leu
203      65      70      75      80
205 Asn Val Ala Glu Ala Ser Phe Leu Ala Ser Met Val Lys Asn Thr Gly
206      85      90      95
208 Asn Ser Tyr Gln Tyr Ile Trp Ile Gly Leu His Asp Pro Thr Leu Gly

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/717,095

DATE: 02/11/2003

TIME: 20:11:15

Input Set : N:\Crif3\RULE60\09717095.txt

Output Set: N:\CRF4\02112003\I717095.raw

```

209           100           105           110
211 Gly Glu Pro Asn Gly Gly Gly Trp Glu Trp Ser Asn Asn Asp Ile Met
212           115           120           125
214 Asn Tyr Val Asn Trp Glu Arg Asn Pro Ser Thr Ala Leu Asp Arg Gly
215           130           135           140
217 Phe Cys Gly Ser Leu Ser Arg Ser Ser Gly Phe Leu Arg Trp Arg Asp
218           145           150           155           160
220 Thr Thr Cys Glu Val Lys Leu Pro Tyr Val Cys Lys Phe Thr Gly
221           165           170           175

```

223 (2) INFORMATION FOR SEQ ID NO: 4:

225 (i) SEQUENCE CHARACTERISTICS:

226 (A) LENGTH: 175 amino acids

227 (B) TYPE: amino acid

228 (D) TOPOLOGY: linear

230 (ii) MOLECULE TYPE: protein

232 (vi) ORIGINAL SOURCE:

233 (A) ORGANISM: Homo sapiens

236 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

238 Met Leu Pro Pro Met Ala Leu Pro Ser Val Ser Trp Met Leu Leu Ser
239 1           5           10           15
241 Cys Leu Met Leu Leu Ser Gln Val Gln Gly Glu Glu Pro Gln Arg Glu
242           20           25           30
244 Leu Pro Ser Ala Arg Ile Arg Cys Pro Lys Gly Ser Lys Ala Tyr Gly
245           35           40           45
247 Ser His Cys Tyr Ala Leu Phe Leu Ser Pro Lys Ser Trp Thr Asp Ala
248           50           55           60
250 Asp Leu Ala Cys Gln Lys Arg Pro Ser Gly Asn Leu Val Ser Val Leu
251           65           70           75           80
253 Ser Gly Ala Glu Gly Ser Phe Val Ser Ser Leu Val Lys Ser Ile Gly
254           85           90           95
256 Asn Ser Tyr Ser Tyr Val Trp Ile Gly Leu His Asp Pro Thr Gln Gly
257           100          105          110
259 Thr Glu Pro Asn Gly Glu Gly Trp Glu Trp Ser Ser Ser Asp Val Met
260           115          120          125
262 Asn Tyr Phe Ala Trp Glu Arg Asn Pro Ser Thr Ile Ser Ser Pro Gly
263           130          135          140
265 His Cys Ala Ser Leu Ser Arg Ser Thr Ala Phe Leu Arg Trp Lys Asp
266           145          150          155          160
268 Tyr Asn Cys Asn Val Arg Leu Pro Tyr Val Cys Lys Phe Thr Asp
269           165          170          175

```

271 (2) INFORMATION FOR SEQ ID NO: 5:

273 (i) SEQUENCE CHARACTERISTICS:

274 (A) LENGTH: 174 amino acids

275 (B) TYPE: amino acid

276 (D) TOPOLOGY: linear

278 (ii) MOLECULE TYPE: protein

280 (vi) ORIGINAL SOURCE:

281 (A) ORGANISM: Rattus rattus

284 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

RAW SEQUENCE LISTING

DATE: 02/11/2003

PATENT APPLICATION: US/09/717,095

TIME: 20:11:15

Input Set : N:\Crif3\RULE60\09717095.txt

Output Set: N:\CRF4\02112003\I717095.raw

```

286 Met Leu Pro Arg Val Ala Leu Thr Thr Met Ser Trp Met Leu Leu Ser
287 1 5 10 15
289 Ser Leu Met Leu Leu Ser Gln Val Gln Gly Glu Asp Ala Lys Glu Asp
290 20 25 30
292 Val Pro Thr Ser Arg Ile Ser Cys Pro Lys Gly Ser Arg Ala Tyr Gly
293 35 40 45
295 Ser Tyr Cys Tyr Ala Leu Phe Ser Val Ser Lys Ser Trp Phe Asp Ala
296 50 55 60
298 Asp Leu Ala Cys Gln Lys Arg Pro Ser Gly His Leu Val Ser Val Leu
299 65 70 75 80
301 Ser Gly Ser Glu Ala Ser Phe Val Ser Ser Leu Ile Lys Ser Ser Gly
302 85 90 95
304 Asn Ser Gly Gln Asn Val Trp Ile Gly Leu His Asp Pro Thr Leu Gly
305 100 105 110
307 Gln Glu Pro Asn Arg Gly Gly Trp Glu Trp Ser Asn Ala Asp Val Met
308 115 120 125
310 Asn Tyr Phe Asn Trp Glu Thr Asn Pro Ser Ser Val Ser Gly Ser His
311 130 135 140
313 Cys Gly Thr Leu Thr Arg Ala Ser Gly Phe Leu Arg Trp Arg Glu Asn
314 145 150 155 160
316 Asn Cys Ile Ser Glu Leu Pro Tyr Val Cys Lys Phe Lys Ala
317 165 170

```

319 (2) INFORMATION FOR SEQ ID NO: 6:

321 (i) SEQUENCE CHARACTERISTICS:

322 (A) LENGTH: 174 amino acids

323 (B) TYPE: amino acid

324 (D) TOPOLOGY: linear

326 (ii) MOLECULE TYPE: protein

328 (vi) ORIGINAL SOURCE:

329 (A) ORGANISM: Rattus rattus

332 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```

334 Met Leu Pro Arg Leu Ser Phe Asn Asn Val Ser Trp Thr Leu Leu Tyr
335 1 5 10 15
337 Tyr Leu Phe Ile Phe Gln Val Arg Gly Glu Asp Ser Gln Lys Ala Val
338 20 25 30
340 Pro Ser Thr Arg Thr Ser Cys Pro Met Gly Ser Lys Ala Tyr Arg Ser
341 35 40 45
343 Tyr Cys Tyr Thr Leu Val Thr Leu Lys Ser Trp Phe Gln Ala Asp
344 50 55 60
346 Leu Ala Cys Gln Lys Arg Pro Ser Gly His Leu Val Ser Ile Leu Ser
347 65 70 75 80
349 Gly Gly Glu Ala Ser Phe Val Ser Ser Leu Val Thr Gly Arg Val Asn
350 85 90 95
352 Asn Asn Gln Asp Ile Trp Ile Trp Leu His Asp Pro Thr Met Gly Gln
353 100 105 110
355 Gln Pro Asn Gly Gly Gly Trp Glu Trp Ser Asn Ser Asp Val Leu Asn
356 115 120 125
358 Tyr Leu Asn Trp Asp Gly Asp Pro Ser Ser Thr Val Asn Arg Gly Asn
359 130 135 140

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/717,095

DATE: 02/11/2003

TIME: 20:11:16

Input Set : N:\Crf3\RULE60\09717095.txt

Output Set: N:\CRF4\02112003\I717095.raw

L:32 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:33 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]